Modeling of Compressive Strength of Admixture-based Self Compacting Concrete using Fuzzy Logic and Artificial Neural Networks

Venu Malagavelli and Prasanth Abraham Manalel
Department of Civil Engineering, BITS-Pilani, Hyderabad Campus, Hyderabad, Andhra Pradesh, 500078, India

Corresponding Author: Venu Malagavelli, Department of Civil Engineering, BITS-Pilani, Hyderabad Campus, Jawahar Nagar, Shameerpet Mandal, Ranga Reddy Dist., Hyderabad, Andhra Pradesh, 500078, India
Tel: +91 98662 71387

ABSTRACT
In the recent past, the applications of artificial intelligence in data analysis and prediction has indisputably increased. Fuzzy Logic (FL) and Artificial Neural Networks (ANN) find extensive applicability with the aim of achieving human-like or superior performance. FL is used in the fields of consumer products, industrial process control, medical instrumentation and portfolio selection while ANN applications include system identification and control, decision making, pattern recognition, sequence recognition, visualization, data mining and financial applications. FL and ANN have the ability to learn from its environment and to improve its performance through learning. This study presents application of FL and ANN in predicting the 28 day compressive strength of self-compacting concrete containing mineral and chemical admixtures. This becomes extremely advantageous in predicting the compressive strength of Self Compacting Concrete (SCC) mixes containing binary, ternary or quaternary blends. The results obtained from the fuzzy logic prediction model and from ANN training, testing and validation were compared with the experimental values from the literature and its performance was evaluated in terms of root mean square error, correlation coefficient, coefficient of performance, mean absolute error and percentage mean relative error.

Key words: Fuzzy logic, compressive strength, mamdani, artificial neural networks, mineral admixture, Levenberg-marquardt, prediction

INTRODUCTION
Concrete is a heterogeneous building material extensively used in the construction industry. This is due to its vast applications, availability of raw materials, economy, ability to take up any form, ability to be fabricated practically anywhere and its inherent durability. Numerous structures and historical landmarks advocate the advantages of concrete durability and versatility. The 90’s witnessed the conceptualization of an innovative concrete known as Self Compacting Concrete (SCC). Since then, SCC has gained popularity due its inherent features (Okamura and Ouchi, 2003). It flows due to its self-weight, occupies the formwork and attains full compaction. Hence, vibration is eliminated which improves productivity and working conditions. It also possesses higher quality, performance and improves working conditions and gives better surface finishes. It can be...
used in areas of congested reinforcement. Using SCC results in faster construction, it offers placement of concrete rapidly. Advantages of high level of homogeneity, minimal concrete voids, uniform concrete strength and durability is ensured by the fluidity and segregation resistance of SCC. SCC is often produced with low water cement ratio providing the prospect of high early strength, earlier stripping of formwork and faster use of elements and structures (Murthy et al., 2012). Many researchers have suggested different mix design procedures. But, there is no single standard method recognized for SCC mix design. Many academic organizations, ready-mixed, precast and contracting establishments have developed their own mix proportioning methods.

SCC is characterized by the use of high volumes of cement and chemical admixtures which has increased production costs (Sukumar et al., 2008). A sustainable solution to reduce the cost of SCC is by using fillers or mineral admixtures such as limestone powder, fly ash, GGBS, etc., which are finely divided materials added to concrete during mixture procedure. When significant percentage of cement is replaced by such mineral admixtures, the cost of production can be reduced. Incorporation of mineral admixtures also results in minimal use of viscosity-enhancing chemical admixtures as they improve rheological and durability properties. In addition to the mentioned, it also reduces thermally-induced cracking of concrete due to the reduction in the overall heat of hydration and helps improve the workability and long-term properties of concrete (Park et al., 2005; Uysal and Yilmaz, 2011; Aldea et al., 2000).

Researches have been conducted in studying the properties of concrete containing these mineral admixtures. Also, the optimized percentage of these supplementary materials has been established. But, these researches have been carried out for different requirements and applications at varying conditions. Due to difference in quality and quantity of material constituents and depending on the design specifications adopted, the SCC mix from one source may vary from another. This creates an uncertainty in establishing a general relationship between mineral admixtures and cement ratio, chemical admixtures and w/c ratio and the like.

This study presents application of FL and ANN in evaluating the 28 day compressive strength of concrete containing mineral and chemical admixtures. This is done by inputting mixes and their respective strengths from literature into FL and ANN which creates a relationship between the input and output parameters which can be either linear or non-linear. In FL, this is taken care of by use of triangular and Gaussian membership functions, while it is assumed as non-linear represented by the Tan-Sigmoid Transfer function. This method is extremely useful in designing mixes for binary, tertiary and quaternary mixes, where combination of cement and number of mineral admixtures are used in the same mix.

FUZZY LOGIC THEORY

Zadeh (1965) introduced the concept of fuzzy logic substituting the Aristotelian logic which has two definite and distinct possibilities only, i.e., 1 or 0 (Aldea et al., 2000). FL provides a mathematical framework to deal imprecision associated with the description of an attribute. This is accounted to the absence of sharply defined criteria. Uncertainties do not mean arbitrary, probabilistic and stochastic deviations, all of which are based on the numerical data. Fuzzy set theory provides an orderly calculus to deal with such information. Fuzzy approach performs numerical computation by using linguistic labels stimulated by membership functions. Therefore, Zadeh introduced linguistic variables as variables whose values are sentences in a natural or artificial language. Fuzzy concepts and systems attracted attention after a real control application in 1975 conducted by Mamdani and Assilian (1975). In real-time situations, many sets have more
Fig. 1: Typical membership function for different category

than an either-or principle for membership. The key idea in FL is the allowance of partial belongings of any object to different subsets of the universal set instead of belonging to a single set totally. Partial belonging to set can be described numerically by a membership function which assumes values between 0 and 1. For instance, Fig. 1 shows a typical membership function for small, medium and large class sizes in a universal set ‘u’. Hence, these verbal assignments are fuzzy subsets of the universal set. In Fig. 1, set values less than 2 are definitely ‘small’, those between 4 and 6 are certainly ‘medium’, while values larger than 8 are definitely ‘large’. However, intermediate values such as 2.2 partially belong to the subsets ‘small’ and ‘medium’. In fuzzy terminology 2.2 has a membership value of 0.9 in ‘small’ and 0.1 in ‘medium’ but 0.0 in ‘large’ subsets. Thus, the transition from membership to non-membership is gradual rather than abrupt (Mamdani and Assilian, 1975; Demir, 2005).

A general Fuzzy Inference System (FIS) has basically four components, fuzzification, fuzzy rule base, fuzzy output engine and defuzzification. Fuzzification converts each piece of input data to degrees of membership by a lookup in one or more several membership functions. Fuzzy rule base contains rules that include all possible fuzzy relation between inputs and outputs. These rules are expressed in the if-then format. There are basically two types of rule base: Sugeno type and Mamdani type. Fuzzy inference engine takes into consideration all the fuzzy rules in the fuzzy rule base and learns how to transform a set of inputs to corresponding outputs. There are basically two kinds of inference operators, minimization (min) and product (prod). Defuzzification converts the resulting fuzzy outputs from the fuzzy inference engine to a number. There are mainly 5 in-built defuzzification methods namely centroid, bisector, middle of maximum, smallest of maximum and largest of maximum. In the present study, the fuzzy model used is of Mamdani fuzzy rule type and the prod method was employed because of its more precise result methodology (McNeill and Thro, 1994). The AND relation was used between the variables. In the present study, Gaussian membership functions are used which are characterized by 2 parameters, mean and standard deviation. Gaussian membership functions are becoming increasingly popular for specifying fuzzy sets due to their smoothness. Gaussian functions are well known in probability and statistics and they possess useful properties such as in variance under multiplication and Fourier transform. For defuzzification, centroid method has been applied.

ARTIFICIAL NEURAL NETWORK THEORY

Inspired by the biological functioning of the human brain, ANN was developed to mimic its essential features in correlating, decision making and pattern recognition. Neural networks are
networks of many simple elements known as neurons, operating in dense parallel interconnections. The structure of a biological neuron consists of a central cell body, an axon and a multilayer of dendrites (Tapkin et al., 2010). The output from the cell is transmitted through the axon to synapses where the outputs are increased or decreased depending upon the synaptic weights before passing it to the ANN models. These have the ability to learn and generalize the problems even when input data contains error or is incomplete. This uncertainty is accounted for during the training process. Activation functions are used in communicating outputs to neurons which is received in the form of weighted inputs. Thus, information is relayed in the form of massive cross-weighted interconnections (Gupta, 2013). In this case, the single layer neural network is used, that takes input from the outside of the networks and transmits their output to the outside of the networks, otherwise, the neural networks are considered multi layered. There are mainly three basic steps in neural network, network training, testing and validation and execution. The weights of the neural network are adjusted through the training process while the training stage is referred to as learning. It involves modification of connection weights by means of a learning rule. The learning process is done by assigning weights and biases computed from a set of training data or by adjusting the weights according to a certain condition. In other words, a neural network generalizes a relationship between inputs and outputs based on the learning data which would have some agreement with the testing data. The initial weights and biases joining nodes of an input layer, hidden layers and an output layer are commonly assigned randomly. It can also be assigned manually if the system is well understood. As input data are passed through hidden layers and sigmoidal activation functions are generally used (Anderson, 1983; Gunaydin and Dogan, 2004). This is because the function is continuous and it has a very simple derivative that is useful for development of learning algorithms and also it represents the processing of biological neuron. Linear activations functions can be used depending on the relationship between the input and output [11 mathworks]. During the training procedure, new weights and biases can be repeatedly obtained until there are no errors. Ultimately, learning corresponds to determining the weights and biases associated with the connections in the networks. The back propagation networks were used in this study. The term back propagation refers to the manner in which the gradient is computed for nonlinear multilayer networks. The default back propagation training algorithm is Levenberg-Marquardt (trainlm). The learning mechanism of the back propagation networks is a generalized delta rule that performs a gradient descent on the error space. Modification in the weights is done by minimizing the error between the actual and predicted results by performing the least mean square procedure (Bilgehan and Turgut, 2010).

for modeling of $f_c$ of cement mortar using the common three layer feed-forward type of artificial neural networks. Mukherjee and Biswas (1997) constructed a model with two hidden layers feed-forward type of ANN for predicting the mechanical behavior of concrete at high temperature.

**MATERIALS AND METHOD**

In case of random data, successful prediction by FL and ANN can be obtained only if the data set is large enough and the parameters selected to represent the data set is apt. Thus, availability of large variety of experimental data was required to develop the relationship between the mixture variables of SCC and its measured properties. For this purpose, a database of 169 mixes from the literature (Sukumar et al., 2008; Grbic et al., 2008; Brouwers and Radix, 2005; Fathi et al., 2013; Baskar et al., 2012; Sahmaran et al., 2009; Sifkas and Trezos, 2013; Memon et al., 2011; Gesoglu and Ozbay 2007; Gesoglu et al., 2009; Rahman et al., 2014; Siddique, 2011; Valcuende et al., 2012; Felekoglu et al., 2007; Sonebi, 2004; Phani et al., 2013; Rao et al., 2009; Gettu et al., 2002; Gandage et al., 2013) was retrieved having mixture composition with comparable physical and chemical properties. The omission of one or more of SCC properties in some studies and the uncertainty of mixture proportions and testing methods in others was responsible for setting the condition for identification of data. The Fuzzy Logic System (FLS) as shown in Fig. 2, implements a nonlinear mapping between its inputs and outputs. Experimental variables consist of Cement (C), Limestone Powder (LP), Fly Ash (FA), Ground Granulated Blast Furnace Slag (GGBS), Silica Fume (SF), Rice Husk Ash (RHA), Coarse Aggregate (CA), Fine Aggregate (F), water, Super Plasticizer (SP), Viscosity Modifying Agent (VMA) as inputs and 28 day Compressive Strength (CS) as output. Table 1 shows the range of the inputs and output.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Parameter</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>C</td>
<td>150.0</td>
<td>570.0</td>
</tr>
<tr>
<td></td>
<td>LP</td>
<td>0.0</td>
<td>272.0</td>
</tr>
<tr>
<td></td>
<td>FA</td>
<td>0.0</td>
<td>350.0</td>
</tr>
<tr>
<td></td>
<td>GGBS</td>
<td>0.0</td>
<td>330.0</td>
</tr>
<tr>
<td></td>
<td>SF</td>
<td>0.0</td>
<td>250.0</td>
</tr>
<tr>
<td></td>
<td>RHA</td>
<td>0.0</td>
<td>200.0</td>
</tr>
<tr>
<td></td>
<td>CA</td>
<td>500.0</td>
<td>927.0</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>478.0</td>
<td>1135.0</td>
</tr>
<tr>
<td></td>
<td>Water</td>
<td>94.5</td>
<td>250.0</td>
</tr>
<tr>
<td></td>
<td>SP</td>
<td>0.0</td>
<td>22.50</td>
</tr>
<tr>
<td></td>
<td>VMA</td>
<td>0.0</td>
<td>1.23</td>
</tr>
<tr>
<td>Output</td>
<td>CS</td>
<td>10.2</td>
<td>117.03</td>
</tr>
</tbody>
</table>

**Fig. 2:** Schematic representation of a fuzzy logic system
For the 169 set of mixes, Mamdani-based FIS was generated using both Gaussian and Triangular membership functions separately. Clustering of data was done and membership functions were created to represent each of the cluster. The functions have been chosen such that all the values in the range are considered. For creation of Gaussian fuzzy set, mean and standard deviation of the each respective cluster is required while the triangular fuzzy set requires end points and mean. For example, consider the Gaussian membership function named ‘475’ for input ‘cement’. This membership function was generated by taking a set of values having mean close to 475 and considering the standard deviation of the cluster.

These parameters are then used to generate the membership function. Similarly, all the other membership functions for each of the input and output parameters were generated. Block diagram representing fuzzy logic modeling of input and output parameters are shown in Fig. 3. Once modeling is done, rules are created which can be viewed in the rule viewer.

![Diagram](image)

![Graph](image)

Fig. 3(a-l): Continue
Fig. 3(a-l): Continue
Fig. 3(a-l): Membership function for (a) Cement, (b) Limestone powder, (c) Fly ash, (d) GGBS, (e) Silica fume, (f) RHA, (g) Coarse aggregate, (h) Fine aggregate, (i) Super plasticizer, (j) Water, (k) VMA and (l) 28 Day Compressive strength
Fig. 4: Typical architecture of a neural network

Different ANN architectures, shown in Fig. 4, were tried and then the appropriate model structure was determined for the data sets. Numerous trials were carried out in the neural network environment to determine optimum neuron number of the hidden layers. In the learning stage, the feed-forward back propagation type network was used. The training function, Levenberg-Marquardt (trainlm) and default performance function Mean Square Error (MSE) were employed.

Throughout all ANN simulations, the learning rates were used for increasing the convergence velocity. The tangent sigmoid and linear functions were used for the activation functions of the hidden and output nodes, respectively. The training phase was stopped after 5000 epochs when the variation of error became sufficiently small.

Gradient descent algorithm back-propagation learning rule was employed with activation functions as tangent sigmoid (tansig). Learning rate was considered as 0.3 with training performance goal $10^{-5}$, momentum constant 0.9 and maximum number of epochs 5000. After carrying out numerous trainings in the neural network simulation, the optimum hidden neuron number and hidden layer number were determined as 20 and 1, respectively.

RESULTS AND DISCUSSION

The data set is input into fuzzy logic and rule sets are generated. Once all the rules are generated based on the input, rule viewer can be generated which gives us output based on a combination of input values. Any combination of inputs can be evaluated and the output can be predicted.

The data set is input into ANN of which 70% is used for learning, 15% for testing and 15% for validation. Here, parametric study involves the optimization of Learning Rate (LR) and hidden neuron number. The learning rate is to determine the changes to the weights and biases. The performance of the algorithm is sensitive to the proper selection of the learning rate. If the learning rate is made too large, the algorithm becomes unstable and oscillates. If the learning rate is set too small, the algorithm takes a long time to converge.
Fig. 5: R-values versus different hidden neuron numbers for the data set.

The accuracy of these models can be evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), correlation coefficient (R), coefficient of determination ($R^2$) and Percentage Mean Relative Error (PMRE):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\text{Actual} - \text{predicted})^2}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |\text{Predicted} - \text{actual}|$$

$$R^2 = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2}$$

$$\text{PMRE} = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{y - \hat{y}}{y} \right|$$

where, y is the experimental value, $\hat{y}$ is the predicted value of y and $\bar{y}$ is the mean of the experimental values.

Figure 5 and Table 2 show the R values for different hidden neuron numbers. It can be seen that the highest R value is obtained when single layer containing 20 hidden neurons is used. The optimum learning rate is found to be 0.3 for the data set as presented in Table 3 and Fig. 6.

Error analysis results reveal the accuracy of the model. Coefficient of correlation (R) for the Gaussian FL model is obtained as 0.98 which is the best model for the data set considered. This model is apt and can be used in practice considering $R = 1$ is the best fit. The other models are also reasonably useful as these are approximate methods.

After rules are created, surfaces can be generated. Figure 7-12 gives surface plots of cement-mineral admixture 28 day compressive strength and cement water 28 day compressive strength relationships through the plots. The 3-D shaded plot containing 3 parameters can be generated. These plots give the relationship between cement a mineral admixtures and compressive strength. These surface diagrams are a function of the input set.
Fig. 6: R-values versus learning rate for the data set

Table 2: Performances of the network architecture for different hidden neuron numbers

<table>
<thead>
<tr>
<th>Hidden neuron No.</th>
<th>Learning rate</th>
<th>Epoch No.</th>
<th>R-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.3</td>
<td>5000</td>
<td>0.8928</td>
</tr>
<tr>
<td>10</td>
<td>0.3</td>
<td>5000</td>
<td>0.9352</td>
</tr>
<tr>
<td>15</td>
<td>0.3</td>
<td>5000</td>
<td>0.9412</td>
</tr>
<tr>
<td>20</td>
<td>0.3</td>
<td>5000</td>
<td>0.9669</td>
</tr>
<tr>
<td>25</td>
<td>0.3</td>
<td>5000</td>
<td>0.9624</td>
</tr>
<tr>
<td>30</td>
<td>0.3</td>
<td>5000</td>
<td>0.9467</td>
</tr>
<tr>
<td>35</td>
<td>0.3</td>
<td>5000</td>
<td>0.9464</td>
</tr>
<tr>
<td>50</td>
<td>0.3</td>
<td>5000</td>
<td>0.9463</td>
</tr>
<tr>
<td>100</td>
<td>0.3</td>
<td>5000</td>
<td>0.9183</td>
</tr>
<tr>
<td>150</td>
<td>0.3</td>
<td>5000</td>
<td>0.8813</td>
</tr>
<tr>
<td>200</td>
<td>0.3</td>
<td>5000</td>
<td>0.7950</td>
</tr>
</tbody>
</table>

Table 3: Performances of the network architecture for different learning rate

<table>
<thead>
<tr>
<th>Hidden neuron No.</th>
<th>Learning rate</th>
<th>Epoch No.</th>
<th>R-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.1</td>
<td>5000</td>
<td>0.9399</td>
</tr>
<tr>
<td>20</td>
<td>0.2</td>
<td>5000</td>
<td>0.9677</td>
</tr>
<tr>
<td>20</td>
<td>0.3</td>
<td>5000</td>
<td>0.9699</td>
</tr>
<tr>
<td>20</td>
<td>0.4</td>
<td>5000</td>
<td>0.9553</td>
</tr>
<tr>
<td>20</td>
<td>0.5</td>
<td>5000</td>
<td>0.9420</td>
</tr>
<tr>
<td>20</td>
<td>0.6</td>
<td>5000</td>
<td>0.9274</td>
</tr>
</tbody>
</table>

Table 4: Results of the error analysis parameters

<table>
<thead>
<tr>
<th>Models</th>
<th>RMSE</th>
<th>MAE</th>
<th>$R^2$</th>
<th>$R$</th>
<th>PMRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FL (Gaussian)</td>
<td>3.336</td>
<td>0.121</td>
<td>0.960</td>
<td>0.980</td>
<td>6.206</td>
</tr>
<tr>
<td>FL (Triangular)</td>
<td>5.823</td>
<td>0.452</td>
<td>0.908</td>
<td>0.983</td>
<td>7.957</td>
</tr>
<tr>
<td>ANN</td>
<td>4.589</td>
<td>3.418</td>
<td>0.940</td>
<td>0.970</td>
<td>7.273</td>
</tr>
</tbody>
</table>

Plot of experimental and predicted values from FL (Gaussian and Triangular) and ANN is shown in Fig. 13. The error analysis parameters for each of the model considered is listed in Table 4. From the error analysis fuzzy logic using Gaussian is giving better results i.e., closer to the experimental values compared to others i.e., ANN and fuzzy logic with triangular functions.
Fig. 7: Surface diagram of cement, LP and compressive strength

Fig. 8: Surface diagram of cement, FA and compressive strength

Fig. 9: Surface diagram of cement, GGBS and compressive strength

547
Fig. 10: Surface diagram of cement, SF and compressive strength

Fig. 11: Surface diagram of cement, RHA and compressive strength

Fig. 12: Surface diagram of cement, water and compressive strength
CONCLUSION

FL and ANN models were successful in predicting the 28 day compressive strength for the set of mixes obtained from literature. This is evident from the results of the error analysis. This study indicates the ability of the Mamdani-based model and multilayer feed forward back propagation neural network as good techniques in modeling the admixture-based SCC in the absence of any pre-defined criteria and calculations and performs sufficiently in the estimation of 28 day compressive strength. The Mamdani-based model used both Gaussian and Triangular membership functions and it is seen that using the former gave better results for the data used. Gradient descent algorithm and one hidden layer were employed in the ANN analysis. The advantage of the proposed approach is that the compressive strength for different mix proportions can be corroborated depending on application can be predicted without lengthy trial-and-error experiments thus decreasing material wastes and production costs. Hence, without conducting the experiments, it is possible to predict the 28 day compressive strength of SCC mix. This tool can be utilized in every field where a direct relationship between the input and output variables are absent.

REFERENCES


